Discovering indicators of successful collaboration using tense: automated extraction of patterns in discourse

Abstract
This paper describes a technique for locating indicators of success within the data collected from complex learning environments, proposing an application of e-research to access learner processes and measure and track group progress. The technique combines automated extraction of tense and modality via parts-of-speech (PoS) tagging with a visualization of the timing and speaker for each utterance developed to code and analyze learner discourse, exploiting the results of previous, non-automated analyses for validation. The work is developed using a dataset of interactions within a multi-user virtual environment, and extended to a more complex dataset of synchronous chat texts during a collaborative technology design task. This methodology extends natural language processing into computer-based collaboration contexts, discovering the linguistic micro-events that construct the larger phases of successful design-based learning.

Key Words: discourse analysis, computer supported collaborative learning, complex educational datasets, natural language processing

Introduction

Increasingly complex datasets can be developed from collaborative content creation tasks, combining video, audio, screen capture, and physical and digital artefacts, including a history of the creative process and collaborative interactions, as well as the learners' final digital object/s. These multiple streams of data, subjected to technology-enhanced analytical methods, can inform our understanding of the relationship between the designable elements of a task and the behaviour of the learners. This paper presents an eResearch technique that has the potential to become a part of the canon of applied computational linguistics techniques used in education (Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012) to assist educators in the orchestration and assessment of online collaborative work.

The technology-enhanced analysis technique described in this paper is concerned with the identification of patterns in discourse that indicate successful and unsuccessful collaboration. A background on verb use in the context of collaboration is provided, followed by a description of the methods of automated extraction of this data, the visualizations used, and results and discussion of the linguistic analysis and its implications for our understanding of group development and the design of collaborative learning environments. In this way the work looks towards circumstances in which automated analysis of collaborative learning can assist educators, learners and researchers.

Our previous work (e.g. Kennedy-Clark & Thompson, 2013a; Thompson, Kennedy-Clark, Kelly & Wheeler, 2013) and other work in the area of multimodal learning analytics (Bilkstein, 2013) led us to imagine what would be possible if a teacher was able to "see" computer-mediated collaborative activity while it was in progress. Teachers, armed with an accurate picture of a group's collaboration, could provide targeted feedback and learning support, using indicators from the discourse to aid the orchestration of learning experiences according to the needs of learners in close-to-real time.

The method of automatic data extraction presented here was developed using a dataset of transcribed conversation sequences from a virtual inquiry in a scenario-based multi-user virtual

environment (MUVE). The method was then tested using the synchronous chat data generated by learners engaged in a month-long collaborative design task. We argue that this technique, whilst limited here by small sample sizes, might allow for the identification of the types of mini-event that learners enact and the determination of whether learners complete the transactional functions crucial to task success (e.g. reporting, determining rules, planning, implementing or resolving).

Background

Each decision a researcher makes about the use of one method implicates the use of a learning theory, often to the exclusion of others. The more holistic approach often offered by eResearch methods, and the method offered in this paper, provides a way to interpret and understand student interactions so that we can better plan learning activities and assessment to support student and group achievement of learning outcomes. An automated analysis of student conversations that extracts modality and the tenses of finite clauses provides clear landmarks (for the researcher or instructor) within the varied and complex datasets generated during learner interactions.

As members of a group interact across the time and events related to the task, they continually negotiate and redefine roles, group expectations and group goals that move the group toward the desired outcome (Gee & Green, 1998, Eggins & Slade, 1997). Within these groups, members are presented with learning opportunities and the ability to have agency in their learning (Gee & Green, 1998). An automated analysis of the moment-by-moment interactions in this context can construct the chain of collaboration while the task is still incomplete, and the specificity of the analysis can be used to guide student learning. It is an alternative to assessing the group's success through examining the outcome of the task, and without the lag associated with a non-automated analysis.

A study of mini-events in CSCL can indicate the type of event a speaker and receiver (person being spoken to) are engaged in. These interactional events include task-related undertakings, such as a *planning* event or a *recount* event and are linked to the tense of the verbal phrase. In English, the tense of the verbal phrase orients the speaker/writer in time (future, past or present) relative to the context of that utterance and to other propositions in the discourse (Jackson & Stockwell, 2011; Martin & Rose, 2007; Willis, 2004): importantly, this is the speaker's experience of time, rather than some external chronology. The choice of modality (how the speaker presents the likelihood, possibility, or obligatory nature of an event or proposition) for the utterance is similarly from the viewpoint of its speaker (the 'deictic centre', Downing & Locke, 2006, p. 353). While English has many resources to realise time and modality, verbs are at the heart of this realisation, and their usefulness as a tool for digital research is suggested by the predictability of their structure: set rules (Willis, 2004) mean the verbal group is relatively easy to locate within the text, and extract. The use of tense has been used for the computational analysis of team co-operation (Ripoche & Sansonnet, 2006), however there appears to be a gap in the literature in terms of studies that focus on what patterns of tense and modality within utterances are able to reveal about a team's progress.

We sought to outline how, over the duration of a computer-supported collaborative task, the modal verbs and tense used by learners form patterns of progress (events). The concept of "collaborative emergence" (Sawyer, 2005; Sawyer & DeZutter, 2009) draws upon theories of distributed cognition and sociolinguistics and contends that interactions at different levels in a creative task can be conceptualised as a collective social process which requires an improvisational approach within an overarching generic structure of design and development. In previous work we have examined different levels of analysis. Informed by systemic functional

linguistics, “macro” features were analysed, such as planning, orientation, agreement and implementation, as well as “micro” features such as pronoun use (Kennedy-Clark & Thompson, 2011, 2012, 2013b; Thompson et al., 2013). Groups respond differently to the environments that they encounter and that the language they use is as relevant a focus for digital research in CSCL as student engagement with the task and interface. Sawyer (2000) supports this approach:

> When an actor takes a dialogue turn, one possible path is chosen, and many other potential paths are closed off; … it’s hard for us to see that the other paths were ever there at all. The importance of collaborative emergence in dialogue only becomes clear when we spend some time analyzing the many possible paths that the dialogue might have taken at each turn (p.182).

Methods

Two datasets are used in this paper: the Virtual Worlds Dataset (VWD), and the Design Team Dataset (DTD). In both cases previous research has reported on the extraction, analysis and findings from these mini-corpuses of collaborative problem-based learning activities to identify overlaps, indicators of progress through a task, and form new conclusions in our understanding of complex learning environments.

Virtual Worlds Dataset

The VWD was obtained from the recording of dyads’ in-world exchanges during an inquiry task using a MUVE. Dyads participated in a face-to-face environment, and audio and screen recordings were collected. In previous analyses, patterns of interaction within the dyads and with the MUVE (Kennedy-Clark & Thompson, 2013a, 2013b) were identified. The following is based on these findings. Dyad 1 engaged in turn taking, but did not cycle through agreement patterns and enacted few pathways to decision-making: they focused on navigation through the MUVE and not the task or topic; decisions were not task-related; they responded to the ‘here-and-now’ occurrences and did not plan or recount task-related actions. Dyad 6 established a cyclic pattern at the macro levels of their collaboration: they sought confirmation from each other prior to implementation, and task-related actions were a significant behaviour; they extended beyond the here-and-now to plan actions and develop an effective routine to their interactions with the environment and each other. Overall they were successful in achieving a problem solution.

Design team dataset

The DTD was collected over one month, and we will discuss the synchronous, online chat data (2182 utterances) of four postgraduate students participating in a design task. Previous research has concluded that this group underwent a cyclical pattern of decision-making (Reimann, Frerejean & Thompson, 2009) and provided insights into the role of tool use in the complexity of the decision-making processes (Thompson & Kelly, 2012). Recent work has shown that patterns of pronoun use could be used as indicators of transitions between phases of the design work (Thompson et al., 2013).

Linguistically, the first dataset is spoken (researcher-transcribed utterances) and the second is written language, although termed ‘chat’, which misleadingly suggests an oral form (Reeder, Macfadyen, Roche, & Chase, 2004). The similarities of the contexts, however (speaker characteristics and generation, university setting, a collaborative group or pair largely independent of the teacher, performing computer- and task-focussed activities), mean that we can expect the two datasets to be closer in verb usage and temporal phasing than the spoken/written

contrast suggests. The intrusion of the written 'voice' of the instructional manual included in the transcribed VWD utterances also reduces the distance in mode between the probable register-related frequencies of tense choice between the dataset texts.

Parts of Speech (PoS) Tagging

Techniques from computational linguistics have been widely used to automate the analysis and support of collaborative learning (Mu, et al., 2012; Rosé et al., 2008; Trausan-Matu & Rebedea, 2010). Parts-of-speech tagging (PoS) in particular is a technique for automatically "tagging" each word within a corpus of text grammatically, as a part of speech such as a verb, noun or adjective (DeRose, 1988; Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012; Schmid, 1994). This is achieved by creating a statistical profile of words and their part-of-speech (within different contexts) from a very large pre-tagged corpus. PoS tagging has recently been used as an aid to the human analysis of learning data (Rosé et al., 2008; Thompson et al., 2013). We suggest that the relationship between the patterns of verbs that this technique provides and the mini-events of the group could be presented to educators through visualisations such as those exemplified below.

In this work PoS tagging was used to identify: (i) incidences of verb usage; (ii) tense of the verb used; (iii) verb modifiers. While commercial coding packages can be configured to handle live data according to rules through coding nodes, the rapidity and simplicity of a tagger's operation is more than sufficient, free, and purpose-built for this method and objective. In this work a tagger was trained using the Penn Treebank corpus of over 4.5 million words and applied to the utterances in the Virtual Worlds dataset using Python's natural language processing library, the *Natural Language Toolkit* (Bird, Klein, & Loper, 2009). From these the tags relating to verbs could be retrieved resulting in the tag set of different verb categories (Table 1).

Insert Table 1 about here

The outcome from PoS tagging is a list of utterances annotated by the tagger with verb and tense based upon the statistical assumptions of the tagger. In order to extract both tense and modality, we need to present not just the morphologically distinct element of the verbal phrase (Table 1), but also the auxiliary and modal components, and so further additions to the Python script were made to extract any instances of modifiers within the verbal group using the words specified in Table 2.

Insert table 2 about here

An extract that is typical of the output from automated PoS tagging and parsing for modifiers is shown in Table 3.

Insert table 3 about here

A potential problem with using this method is that only past and present verb forms in English provide morphological clues for analysis: other distinctions (for example continuous or

completed actions and reference to future times require auxiliary operators). The '-ed' ending marks this verb as past tense:

'I asked William about the resources' [s 18:48:09, 3 Oct, DTD]

but other operators are needed to show the speaker's reference to a future time:

'I will ask Anne about the graphics not working issue' [s 01:39:56, 26 Oct, DTD]

The distinction between real events (concurrent and prior to the speaker's time) and projected or imagined events (often realised with modal or future constructions, and termed 'irrealis'; Hasan et al. 2007, p. 721) is an important one because learners are using their discussion for several key stages in CSCL:

- *Reporting* on completed actions (using past tense, e.g. 'I turned myself into water' [s46, Dyad 1, VWD)
- *Determining the rules of the environment or task* (expressed as simple present tense, e.g. 'I think we use this thing' [s64, Dyad 1, VWD]
- *Commenting on the actions they are implementing, their progress in the task* (often selecting a continuous present form, e.g. 'I am trying to write "the rainy season"' [s73, Dyad 1, VWD]), and
- *Planning the direction and resourcing of the task*, e.g. 'we will handle how we use the ideas in the design phase' [s 00:37:29, 13 Oct, DTD].

The planning function, directed towards the irrealis, is often realised through modality ('that **might** be one of our tasks' [s 18:21:39, 3 Oct], DTD), particularly where the speaker needs to influence their collaborators, or express an imagined event which better matches the desired outcome ('deontic modality'). A semantic approach that links tense and modality is the 'Event orientation' of a message, how remote the event is from the 'now', that is, concurrent with the speaker's time, or habitual; prior events; non-hypothetical events realised by plans or predictions; and remote hypothetical events or conjecture (Hasan et al. 2007, p. 722). The verbs used can be mapped to this continuum. We have used this scale as a finer focus on our analysis.

Reliability

In PoS analysis the accuracy of automated tagging can vary depending upon the corpus used for training and for analysis. Existing studies show that the accuracy of tagging for each word can be expected to be around 97 per cent (Shen, Satta, & Joshi, 2007; Søgaard, 2010; Toutanova, Klein, Manning, & Singer, 2003). This error is consistent with our findings. After manually testing the results from Dyad 1 (VWD), we found that from the 127 instances of verb use tagged by the automated PoS, three were false positives (2.36%). This is within the expected error range (3%). A limitation is that false negatives were not assessed, and, therefore, total error will be higher. As the expected 3 per cent error may not apply when solely considering verbs a conservative estimate of error around 5 per cent can be adopted by doubling the discovered errors.

Identifying the mini-events

Figure 1 about here

Mini-events were identified that demonstrated the dyads' differing responses to the 'here-and-now'. Figure 1a shows the patterns of verb use for each decision the dyad made. In most decisions, the challenges Dyad 1 experienced in their collaboration (as outlined earlier: their focus on navigation rather than task or topic) were also reflected in the patterns created by their use of tense. Figure 1a shows elements in the simple present (VBZ, VBP) dominating and occurring regularly (1), matched by the frequency of utterances with no verb form at all (such as 'oh', or 'unhh') and utterances with ellipsis of many message elements (2). This is possible in this dataset because the references are to elements and characters in the game environment that are immediate in time and space to the speakers. There is scant use of future or modal forms (3), and the continuous form of the present (1) indicates commentary on current action rather than forecasting or solution description. Dyad 1, as they progressed through the task, attempted to make sense of what they were experiencing, but they did not progress to completion in determining the rules of their environment (indicated by the habitual simple present) and the few occurrences of past tense (VBD) do not report successful completion of stages (4).

Figure 1b shows consistent reference to action for each of the decisions, but with more frequent and regular use of future tense forms to guide the collaboration (5). Dyad 6 spent time at the beginning determining the rules of the environment together (using simple present and continuous present, (6)), but then, crucially, checked their findings with their teacher (7): her use of simple presents (VBP) at this point indicates that she is communicating the rules of success. By contrast, the teacher's guidance to Dyad 1 is in imperative form (8) – hints for action, but not progressing the dyad's understanding of the task.

Results and Discussion

The PoS tagging method was then applied to the DTD. Given the large nature of this dataset, not all patterns will be discussed. The focus is on two instances: the first demonstrates the potential, and the second the limitations, of the proposed method.

Figure 2 about here

Figure 2a shows the end of a phase of choosing the model – how the group should find models to choose from, how the model should be chosen, and which resources should be added. In each case, the decision is made and all in the group agree: there is a clear movement from present tense (a comment on where they are in the task sequence, (1)), to a proposal for a future action (2), imperative and agreement (3). The discussion around the coordination task begins with a planning statement in future tense ('we will do it once we have the model', (4)), recounts on action taken (5), and agreeing on implications for their collaboration (6).

In both the completion of subtasks involved in *choosing a model*, and the beginning of a phase of coordination, an automated analysis of the types of tenses and modalities being used in the computer-mediated communication of the group would provide the researcher or teacher with an indication of the stage that the group had reached in its collaborative design task. With the analysis set at an appropriate granularity or exploiting significant variations from the expected verb choices for the register, the teacher would have the information to make an effective intervention, one comparable to those exemplified from the teacher's interruptions in the VWD.

Figure 2b is from a later stage in the group's collaboration, focusing on two decisions: asking for help; what to add to the model. Planning played an important role (7) and past experiences played a greater role in deciding what to add, than in asking for help (8). This example demonstrates that, while automatic tagging and visualisation can indicate the overall shape of the text, there is still a place for the human analyst. In both decisions, at one point, the present tense functions as future tense (9). Using the logical cues that indicate futurity (e.g. 'saturday to monday is cutting it too fine – our deadline is tuesday isn't it' in asking for help, and 'we add one or more additional variables for learners to experiment with' in adding to the model) and the surrounding context, the present tense is actually used as a planning element for future action. In part, the error is a function of the important omissions that occur in natural speech, such as a modal construction.

This method of eResearch has identified patterns of action that are likely to achieve a successful outcome in collaboration. Whilst no direct classroom application has been provided, the technique has potential to assist educators in orchestration and intervention. In both cases, the participants in the datasets examined are constructing a group text that presents both shared understandings and individual perspectives: that is, they are developing a cultural model (Gee & Green, 1998). Eggins and Slade (1997) explain that the grammatical choices of speakers are not conscious, but both enact and confirm cultural practices. Nevertheless, the teacher or guide, informed by an automated analysis, can provide students with a common language to describe their model and a framework for collaboratively creating content. These patterns of modality and tense choice can be identified, at a local level, to inform the design and conduct of learning activities.

Conclusions

The implications of this method are for the management of collaborative work. In addition to the limitation associated with omissions in natural speech, another consideration is the development required for this to work in face-to-face collaborative environments. However, the value of such a method in online environments is clear, and it holds great promise as a tool for teachers to assess a group's progress. Educators in such environments frequently need to manage a number of groups simultaneously, and a manual analysis of the collaboration is not feasible. Automated analysis, using information such as the technique described could be used to strengthen the evidence-based support to group-work management. The use of this technique currently requires high technical literacy to run the program and interpret the results. Future work will focus on packaging several automated techniques into an online environment so that the utility of the technique in a live online classroom can be assessed, and the specific circumstances under which an educator would gain benefit can be determined. The issue of providing structure to students is not straightforward: automated tools allow a personalized approach to be adopted, providing guidance to learners, as they need it.

In regards to the generalisability of the results, it is highly improbable that questions on learning in computer-mediated contexts can be answered by generalisable strategies and recommendations. Instead these questions may best be answered at a localised context. As the use of teaching analytics becomes more common in education, the key challenge will still be to ask the right questions, and multiple modes of analysis and extracted data will be necessary to understand complex datasets and processes of learning. We have demonstrated the usefulness of extracting data (tense and modality) from a complex dataset, and its application to a complex learning environment is innovative, with implications for both design and eResearch methods.

Ultimately, we argue that eResearch methods should be a catalyst to research that integrates learning and teaching practices with the assessment of student learning, informing the design of learning environments.

Acknowledgements
Part of this work was funded by the Australian Research Council grant FL100100203.

References

Bilkstein, P. (2013). Multimodal Learning Analytics. *LAK 2013*, pp. 102-106

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*: O'Reilly Media, Incorporated.

DeRose, S. J. (1988). Grammatical category disambiguation by statistical optimization. *Computational Linguistics, 14*(1), 31-39.

Downing, A., & Locke, P. (2006). *English grammar: A university course* (2nd ed.). Abingdon, Ox: Routledge.

Eggins, S. & Slade, D. (1997). *Analysing casual conversation*. London: Cassell.

Gee, J.P. & Green, J.L. (1998). Discourse analysis, learning and social practise: A methodological study. *Review of Research in Education*, 23, 119-169/

Hasan, R., Cloran, C., Williams, G. & Lukin, A. (2007). Semantic networks: the description of linguistic meaning in SFL. In R. Hasan, C. Matthiessen & J. Webster (Eds). *Continuing discourse on language: A functional perspective* (Vol. 2). London: Equinox.

Jackson , H., & Stockwell, P. (2011). *An Introduction to the Nature and Functions of Language* (2nd ed.). London: Continuum International Publishing Group.

Kennedy-Clark, S., & Thompson, K. (2011). What Do Students Learn When Collaboratively Using A Computer Game in the Study of Historical Disease Epidemics, and Why? *Games and Culture, 6*(6), 1-25.

Kennedy-Clark, S., & Thompson, K. (2012). *Methods of analysis for identifying patterns of problem solving processes in a computer-supported collaborative environment* Paper presented at The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS), Sydney, 2-6 July.

Kennedy-Clark, S., & Thompson, K. (2013a). Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In L. Hwee Ling & F. Sudaweeks (Eds.), *Innovative Methods and Technologies for Electronic Discourse Analysis*. Pennsylvania: IGI-Global.

Kennedy-Clark, S., & Thompson, K. (2013b). Between the lines: The use of discourse analysis in a virtual inquiry to inform learning design. *International Journal of Virtual and Personal Learning Environments*.

Martin, J. R., & Rose, D. (2007). *Working with Discourse: Meaning Beyond the Clause* (2nd ed.). London: Continuum International.

Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning, 7*(2), 285-305.

Reeder, K., Macfadyen, L., Roche, J., & Chase, M. (2004). Negotiating cultures in cyberspace: participation problems and problematics. *Language Learning and Technology, 8*(2), 88-105.

Reimann, P., Frerejan, J. & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. Paper presented at the CSCL Conference, Rhodes, Greece.

Ripoche, Gabriel, & Sansonnet, Jean-Paul. (2006). Experiences in automating the analysis of linguistic interactions for the study of distributed collectives. *Computer Supported Cooperative Work (CSCW), 15*(2-3), 149-183.

Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., et al. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *3*(3), 237-271.

Sawyer, K. (2000). Improvisational Cultures: Collaborative Emergence and Creativity in Improvisation. *Mind, Culture, and Activity. 7*(3), 180-185.

Sawyer, K. (2005). *Social Emergence: Societies as complex systems*. Cambridge: Cambridge University Press.

Sawyer, K., & DeZutter, S. (2009). Distributed creativity: how collective creations emerge from collaborations. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(2), 8192.

Schmid, H. (1994). *Probabilistic part-of-speech tagging using decision trees.* Paper presented at the Proceedings of international conference on new methods in language processing.

Shen, L., Satta, G., & Joshi, A. (2007). *Guided learning for bidirectional sequence classification.* Paper presented at the Annual Meeting-Association for Computational Linguistics.

Søgaard, A. (2010). *Simple semi-supervised training of part-of-speech taggers.* Paper presented at the Proceedings of the ACL 2010 Conference Short Papers.

Thompson, K., & Kelly, N. (2012). *Processes of decision-making with adaptive combinations of wiki and chat tools*. Paper presented at the The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012).

Thompson, K., Kennedy-Clark, S., Kelly, N., & Wheeler, P. (2013). Using automated and fine-grained analysis of pronoun use as indicators of progress in an online collaborative project. *The 10th Int'l Conference on Computer Supported Collaborative Learning,* Madison, Wisconsin, pp. 486-493.

Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003). *Feature-rich part-of-speech tagging with a cyclic dependency network.* Paper presented at the Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1.

Willis, D. (2004). *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.

Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International Journal of Computer-Supported Collaborative Learning, 7*(2), 285-305. doi: 10.1007/s11412-012-9147-y

Ripoche, G., & Sansonnet, J.-P. (2006). Experiences in automating the analysis of linguistic interactions for the study of distributed collectives. *Computer Supported Cooperative Work (CSCW), 15*(2-3), 149-183.

Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 3*(3), 237-271.

Trausan-Matu, S., & Rebedea, T. (2010). A polyphonic model and system for inter-animation analysis in chat conversations with multiple participants *Computational Linguistics and Intelligent Text Processing* (pp. 354-363): Springer.